

Validation of Landsat 8 high resolution Sea Surface Temperature using surfers

Quinten Vanhellemont^{a,*}, Robert J.W. Brewin^b, Philip J. Bresnahan^c, Tyler Cyronak^d

^a Royal Belgian Institute of Natural Sciences, Operational Directorate Natural Environments, Vautierstraat 29, 1000 Brussels, Belgium

^b Centre for Geography and Environmental Science, College of Life and Environmental Sciences, University of Exeter, Penryn Campus, Cornwall TR10 9FE, UK

^c Department of Earth and Ocean Sciences, University of North Carolina Wilmington, Wilmington, NC 28403-5944, USA

^d Department of Marine and Environmental Sciences, Nova Southeastern University, Dania Beach, FL 33004, USA

ARTICLE INFO

Keywords:

Coastal waters
Nearshore
Surface temperature
Landsat 8
TIRS
Validation

ABSTRACT

Nearshore coastal waters are highly dynamic in both space and time. They can be difficult to sample using conventional methods due to their shallow depth, tidal variability, and the presence of strong currents and breaking waves. High resolution satellite sensors can be used to provide synoptic views of Surface Temperature (ST), but the performance of such ST products in the nearshore zone is poorly understood. Close to the shoreline, the ST pixels can be influenced by mixed composition of water and land, as a result of the sensor's spatial resolution. This can cause thermal adjacency effects due to the highly different diurnal temperature cycles of water bodies and land. Previously, temperature data collected during surfing sessions has been proposed for validation of moderate resolution (1 km pixel size) satellite ST products. In this paper we use surfing temperature data to validate three high resolution (100 m resampled to 30 m pixel size) ST products derived from the Thermal InfraRed Sensor (TIRS) on board Landsat 8 (L8). ST was derived from Collection 1 and 2 Level 1 data (C1L1 and C2L1) using the Thermal Atmospheric Correction Tool (TACT), and was obtained from the standard Collection 2 Level 2 product (USGS C2L2). This study represents one of the first evaluations of the new C2 products, both L1 and L2, released by USGS at the end of 2020. Using automated matchup and image quality control, 88 matchups between L8/TIRS and surfers were identified, distributed across the North-Western semihemisphere. The unbiased Root Mean Squared Difference (uRMSD) between satellite and in situ measurements was generally < 2 K, with warm biases (Mean Average Difference, MAD) of 1.7 K (USGS C2L2), 1.3 K (TACT C1L1) and 0.8 K (TACT C2L1). Large interquartile ranges of ST in 5 × 5 satellite pixels around the matchup location were found for several images, especially for the summer matchups around the Californian coast. By filtering on target stability the number of matchups reduced to 31, which halved the uRMSD across the three methods (to around 1.1K), MAD were much lower, i.e. 1.1 K (USGS C2L2), 0.6 K (TACT C1L1), and 0.2 K (TACT C2L1). The larger biases of the C2L2 product compared to TACT C2L1 are caused as a result of: (1) a lower emissivity value for water targets used in USGS C2L2, and (2) differences in atmospheric parameter retrieval, mainly from differences in upwelling atmospheric radiance and lower atmospheric transmittance retrieved by USGS C2L2. Additionally, tiling artefacts are present in the C2L2 product, which originate from a coarser atmospheric correction process. Overall, the L8/TIRS derived ST product compares well with in situ measurements made while surfing, and we found the best performing ST product for nearshore coastal waters to be the Collection 2 Level 1 data processed with TACT.

1. Introduction

The temperature of Earth's surface drives heat exchange between the surface and atmosphere and has important implications for climate in general. Surface temperature is strongly linked to water availability and water use, evapotranspiration, severe weather, and the growth and metabolic rates of organisms, among other processes. Ocean temperature change is described in the 5th Assessment Report (AR5) of

the Intergovernmental Panel on Climate Change (IPCC) as among the three key drivers of climate change impacting coastal systems (Wong et al., 2014), with profound implications for nearshore coastal environments (Harley et al., 2006). The Global Ocean Observing System and Global Climate Observing System consider Sea Surface Temperature (SST) as an Essential Ocean Variable (<https://www.goosocean.org/eov>,

* Corresponding author.

E-mail address: quinten.vanhellemont@naturalsciences.be (Q. Vanhellemont).

<https://doi.org/10.1016/j.ecss.2021.107650>

Received 19 April 2021; Received in revised form 30 October 2021; Accepted 12 November 2021

Available online 1 December 2021

0272-7714/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

accessed, 2021-04-10) and an Essential Climate Variable (<https://gcos.wmo.int/en/essential-climate-variables/>, accessed 2021-04-10). In the aquatic environment, thermal stratification of the water column impacts nutrient and oxygen cycling and is strongly linked to the rise and collapse of phytoplankton blooms of different species (Jones and Gowen, 1990; Wilhelm and Adrian, 2008; Murphy et al., 2011; Martinez et al., 2011) and in shallow well-mixed waters temperature is a determining factor of phytoplankton productivity (Trombetta et al., 2019). The recent proliferation of (harmful) algal blooms can be linked to eutrophication and increasing temperatures (Davis et al., 2009; O'Neil et al., 2012), while falling sea ice coverage and increased Atlantic water inputs are shifting the phytoplankton bloom dynamics in the warming Arctic (Oziel et al., 2017; Neukermans et al., 2018).

Orbital sensors are the only reasonable way to consistently assess surface temperatures (ST) at local to global scales, and as a result satellite derived ST has found many aquatic and terrestrial applications. Moderate resolution imagery was used to construct global climate records of open ocean ST, e.g. Reynolds SST (Reynolds et al., 2002) and the European Space Agency Climate Change Initiative (SST CCI) (Merchant et al., 2014, 2019) that serve as inputs to climate models and various other applications. Higher resolution imagery from Landsat has been used for mapping of river plumes (Brando et al., 2015) and the thermal effluent from power plants (Ingleton and McMinn, 2012) or treated wastewater outfall pipes (Trinh et al., 2017). In the terrestrial realm, urban heat islands have been studied using thermal satellite data (Yuan and Bauer, 2007; Liu and Zhang, 2011), and the occurrence of drought (Wan et al., 2004) and fire (Kaufman et al., 1998; Giglio et al., 2003) are regularly mapped. Evapotranspiration can be estimated with the aid of remotely-sensed ST, and Landsat in particular has the capability to resolve water use at the field level (Anderson et al., 2012, 2018; Senay et al., 2016). The Copernicus Land Service produces hourly land ST at relatively coarse spatial resolution (5/112° grid cell size) derived from multiple geostationary satellite observations (e.g. Freitas et al., 2013), that can represent the diurnal temperature cycle, and provides opportunities for improving state of the art land surface models (Orth et al., 2017; Nogueira et al., 2020).

Various methods exist for retrieving ST from satellite sensors based on observations either in single or multiple channels in the Thermal InfraRed (TIR). Single channel, single window, and split-window methods tend to show similar performance, but have different requirements for ancillary inputs (Sekertekin and Bonafoni, 2020). Single channel methods need a representative atmospheric profile and a set of radiative transfer simulations to derive the downwelling and upwelling atmospheric radiance and the atmospheric transmittance (Barsi et al., 2003; Cook, 2014). Knowing the target emissivity and the surface emitted radiance allows for the computation of ST. Alternatively, single channel algorithms – in this configuration also called single window or mono-window algorithms – can be based on pre-generated calibration coefficients for ranges of Total Columnar Water Vapour (TCWV). These algorithms may be simpler to deploy and may be more computationally efficient as no radiative transfer code has to be run, and can be readily integrated in cloud computing platforms (Ermida et al., 2020). Split-window algorithms similarly use a set of pre-generated coefficients and two thermal bands to derive ST based on an estimate of TCWV and boundary layer temperature (Wan and Dozier, 1996). For Landsat 8, the use of single band methods using band 10 is preferred, due to stray light issues, especially in band 11 (Barsi et al., 2014; Montanaro et al., 2014a,b; Gerace and Montanaro, 2017), and for compatibility with heritage Landsat sensors (Cook, 2014; Malakar et al., 2018; Sekertekin and Bonafoni, 2020).

As the accuracy of the downstream products depends on the accuracy of the input ST, validation of the satellite derived ST using in situ measurements is of crucial importance, and is done either using matchups with contact thermometers (Cook et al., 2014; Malakar et al., 2018; Vanhellemont, 2020a,b) or infrared radiometers (García-Santos et al., 2018; Malakar et al., 2018; Sekertekin, 2019; Vanhellemont,

2020b; Sekertekin and Bonafoni, 2020), e.g. from the SURFRAD network (Augustine et al., 2000). Due to the rather consistent temperature within a satellite pixel footprint and the well known emissivity of water, water targets are often preferred to the much more variable land targets. As was shown in Vanhellemont (2020a) the water ST derived from Landsat 8 compared well with in situ measurements made at measurement towers and buoys in the Belgian coastal zone (bias of < 0.1 K and scatter of < 0.7 K), but these are located several km from the coast and hence results may not be representative for nearshore waters. Recently, the use of benthic loggers (Brewin et al., 2018) and loggers integrated into sports equipment (Brewin et al., 2015, 2017b, 2020b) has been proposed for validation of nearshore satellite ST. The bulk temperature as measured by these accurate, affordable and easily deployed contact thermometers was found to be ≤ 0.13 K warmer compared to the radiative skin temperature measured by an Infrared Sea surface temperature Autonomous Radiometer (ISAR) (Brewin et al., 2021), rather consistent with the -0.17 K bulk-to-skin difference found in other studies (Donlon et al., 1999).

In the present paper, temperature collected during surfing sessions is used for the validation of high resolution imagery from Landsat 8. ST were obtained from the new Landsat Collection 2 Level 2 products, as well as processed with the Thermal Atmospheric Correction Tool (TACT, Vanhellemont, 2020a,b) using Level 1 data from both Collection 1 and Collection 2. The Collection 2 products provided by USGS contain ST and Surface Reflectance (SR) data, and are produced with their internal Landsat Product Generation System (LPGS) versions 15.3 and later. TACT is an open source processor for deriving ST from Landsat sensors (Landsat 5, 7, and 8), based on the libRadtran (Emde et al., 2016) radiative transfer code, that can be run using various atmospheric profile inputs. It is integrated into ACOLITE and freely available from <https://github.com/acolite/acolite>. The main advantage of TACT over Collection 2 data is the public availability of the code, and the use of a free radiative transfer code. In this study, the performance of these different methods is quantified using in situ matchups. To the best of our knowledge, this work presents a first independent intercomparison of Level 1 thermal data from both collections, as well as an evaluation of the standard Level 2 ST product for retrieving (Water) Surface Temperature.

2. Data and methods

2.1. In situ data

In situ measurements were obtained from: (1) surfing sessions in the southern UK, Ireland, and San Diego, California using Tidbit v2 loggers attached to the surfboard leash (Fig. 1; Brewin et al., 2020a,b), and (2) surfing sessions in various locations using a Smartfin (Fig. 1; Bresnahan et al., 2017), with temperature, motion and GPS data downloaded from the Smartfin website (<https://smartfin.axds.co/>). The leash measurements were processed as described in Brewin et al. (2020b), providing geolocated, median sea surface temperature data, with uncertainties, for each surfing session. The data are freely available through Brewin et al. (2020a)

Smartfin data were processed following a method adapted from Brewin et al. (2020b). Firstly, Smartfin data were downloaded through the Smartfin Application Programming Interface (API, at <https://stage.platforms.axds.co>), with example scripts for accessing these data provided at https://github.com/SUPScientist/Smartfin_data_via_Axiom_API. The API provided a combined data file of motion (accelerometer) and ocean (temperature) data at 1/6 Hz. Motion data in the combined data file were subsampled by the data provider, from their original measurement frequency at 6 Hz, to the temperature measurement frequency. The following steps were taken to remove data that were not collected in the water. The rate of change in temperature was computed using the external temperature sensor (located on the tip of the fin), as was the difference between the internal and external temperature sensor



Fig. 1. Photo of equipment used by surfers to measure water temperature. Tidbit sensors (shaded (white) and unshaded sensor) on leash of surfboard, and Smartfin, with integrated environmental sensor package (surfboard fin closest to the leash). Detailed descriptions of the system are provided by [Brewin et al. \(2015, 2017b, 2020b\)](#).

on the Smartfin. Data were flagged when the rate of change exceeded 0.05 K, and where the temperature difference between internal and external temperature sensors was greater than 0.5 K, or less than −0.8 K. For cases where motion data were available, data were flagged when the pitch angle of the accelerometer (smoothed with a 5-minute filter) fell outside predetermined bounds (i.e., making use of the fact that a surfboard is relatively flat when being surfed in the water). Finally, the first and last 5 min of the temperature data were flagged (i.e., typically when the surfer is entering and exiting the water). All data were visually inspected, and for four surfing sessions (where there were no motion data available), additional measurements were flagged manually, that were clearly not characteristic of measurements collected in the ocean at the time. Having identified measurements when the Smartfin was in the water, the median of the temperature data from the external temperature sensor was computed and taken to be the water temperature for the surfing session. Uncertainties were also approximated for each surf, by taking the square root of the sum of the squared calibration error of the Smartfin (set to 0.05 K, based on laboratory tests in [Brewin et al. \(2020a\)](#)) and the squared median absolute deviation in temperature during each surfing session. Following this processing, and consistent with the leash measurements, geolocated, median sea surface temperature data, with uncertainties, were available for each Smartfin surfing session.

Comparisons between leash-based Tidbit and Smartfin measurements have shown excellent agreement, with no systematic differences (after applying a correction for solar heating in the leash processing) and with a mean absolute difference of around 0.07 K (for 141 surfing sessions, see [Brewin et al. \(2020b\)](#)). Both Smartfin and leash data have also been found to be in very good agreement with independent data collected from benthic and pier measurements ([Brewin et al., 2015, 2017b, 2018, 2020b](#)). The Smartfin has also been evaluated against a standard oceanographic temperature instrument (Seabird SBE38) on an oceanographic voyage through the Atlantic ([Brewin et al., 2021](#)), spanning a gradient in sea surface temperature (SST) of 19 K, and found to be in excellent agreement (mean differences and mean absolute differences of −0.01 and 0.06 K, respectively). For the combined leash and Smartfin data, bulk temperatures were adjusted to skin temperatures using a fixed −0.17 K offset ([Donlon et al., 1999](#)).

2.2. Satellite data

Landsat 8 (L8) has two sensors on board ([Irons et al., 2012](#)), the Operational Land Imager (OLI), a 9-band multispectral instrument covering the visible (VIS), near infrared (NIR), and short-wave infrared (SWIR), and the Thermal Infrared Sensor (TIRS) with two bands in

the thermal infrared at around 10 and 11 μm . The OLI multispectral data is recorded at 30 m spatial resolution in 8 channels, and at 15 m resolution in the panchromatic channel. TIRS data is recorded at 100 metre resolution, and is resampled to the OLI 30 m grid during Level 1 processing and geo-orthorectification. Level 1 (L1) top-of-atmosphere data from L8 is currently available in two collections, Collection 1 (C1) and Collection 2 (C2). Released by the United States Geological Society (USGS) near the end of 2020, C2 for the first time includes standard Level 2 (L2) surface reflectance (SR) and surface temperature (ST) products ([USGS, 2020](#)). For L8/TIRS, the C2L1 processing contains several enhancements compared to C1L1, notably a change in absolute calibration, to reduce radiometric calibration errors after the 2017 stray light correction ([Gerace and Montanaro, 2017](#)), and new relative (detector-to-detector) gains to reduce along track striping that progressively got worse during the mission (<https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-2-level-1-data>, accessed 2021-04-10). C1L1 data were obtained from Google Cloud Services (<https://cloud.google.com/storage/docs/public-datasets/landsat>), and C2L1 and C2L2 data were obtained through the USGS EarthExplorer (<https://earthexplorer.usgs.gov>). L1 TIRS data from both C1 and C2 were processed using TACT ([Vanhellemont, 2020a,b](#)) using atmospheric profiles from the 0.25 degree ERA5 climate model, with radiative transfer results at the 0.25 degree spacing linearly interpolated to the individual 30 m pixels. All three ST (TACT C1L1, TACT C2L1, and USGS C2L2) are single band products derived from band 10 (B10), at around 10 μm . The top-of-atmosphere (TOA) radiance (L_t) in B10 is a measurement of the surface radiance (L_s) and atmospheric up- (L_u) and down-welling (L_d) radiances ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$) transmitted through the atmosphere:

$$L_t = \tau \cdot (\epsilon \cdot L_s + (1 - \epsilon) \cdot L_d) + L_u, \quad (1)$$

where ϵ is the target emissivity, and τ the atmospheric transmittance. TACT (both C1L1 and C2L1) and USGS C2L2 use radiative transfer models to derive L_s from L_t , respectively libRadtran ([Emde et al., 2016](#)) and MODTRAN ([Berk et al., 1999](#)), by deriving the atmospheric parameters (L_u , L_d , τ) based on atmospheric profiles, with details provided in [Vanhellemont \(2020a,b\)](#), [Cook et al. \(2014\)](#), and [Malakar et al. \(2018\)](#). C2L2 uses emissivity values based on the ASTER Global Emissivity Dataset ([Hulley et al., 2015](#)) adjusted by the OLI derived Normalised Difference Vegetation and Snow Indices (NDVI and NDSI) for land pixels, and uses a fixed emissivity of 0.9880 for water pixels. TACT is here configured to use a fixed water emissivity value of 0.9926 for B10 ([Vanhellemont, 2020a](#)). With the atmospheric parameters and target emissivity known, L_s can be inverted from Eq. (1), and is then

converted to temperature (in K):

$$T = \frac{K_2}{\ln\left(\frac{K_1}{L_s} + 1\right)}, \quad (2)$$

where the L8/TIRS B10 specific coefficients are $K_1 = 774.8853 \text{ W m}^{-2} \text{ sr}^{-1} \mu\text{m}$ and $K_2 = 1321.0789 \text{ K}$. For most plots, ST were converted from K to °C (using an offset of -273.15) for familiarity with the encountered temperature ranges.

L1 OLI TOA reflectance (ρ_t) data were processed using ACOLITE (Vanhellemont, 2019, 2020c) to retrieve surface reflectances (ρ_s) using the Dark Spectrum Fitting (DSF) method with a fixed atmospheric path reflectance over a $24 \times 24 \text{ km}$ region of interest centred on the in situ data locations. OLI ρ_t and ρ_s were used for quality control of matchup data (see next section).

2.3. Matchups

Same-day measurements made in situ and by satellite were identified as matchups. A 5×5 pixel box (at the 30 m OLI grid) centred on the pixel containing the in situ coordinate was extracted from the satellite imagery. These boxes were then quality controlled using simple thresholds on the OLI reflectance. Pixels were masked if the top-of-atmosphere reflectance (ρ_t) in any band was greater than 0.3, rejecting land, clouds, high glint and floating objects or vegetation. Further tests were done on the $1.3 \mu\text{m}$ cirrus band, masking pixels where ρ_t ($1.3 \mu\text{m}$) > 0.01 to reject cirrus clouds, and the $1.6 \mu\text{m}$ band, masking pixels where ρ_t ($1.6 \mu\text{m}$) > 0.1 refining the test on floating objects and high glint. Pixels with surface reflectance (ρ_s) greater than 0.15 in any band were also masked, rejecting atypical water spectra.

The matchups have varied fractions of masked pixels (MP, 0 to 24; boxes with 25 MP were rejected) and interquartile range (IQR) of temperature (up to 4K) in the 5×5 pixel extracted box, and the effects of filtering based on these values was further explored. The median value (P50) and interquartile range (IQR, P75–P25) were computed for each 5×5 pixel box to compare with the in situ measurements. Robust statistics were computed, as commonly used in SST validation studies (Merchant and Harris, 1999; Minnett et al., 2019). The Reduced Major Axis (RMA) regression line was computed for the matchups, as were the Mean Average Differences (MAD) and unbiased Root Mean Squared Differences (uRMSD) between in situ and satellite:

$$MAD = \sum_{i=0}^n \frac{y_i - x_i}{n}, \quad (3)$$

$$uRMSD = \sqrt{RMSD^2 - MAD^2}, \quad (4)$$

with RMSD:

$$RMSD = \sqrt{\sum_{i=0}^n \frac{(y_i - x_i)^2}{n}}. \quad (5)$$

3. Results and discussion

3.1. TIRS ST

TIRS ST from the three datasets agree to within a few degrees for the water scenes evaluated here. Among the three, TACT C2L1 data gives the lowest temperatures, and USGS C2L2 the highest. The C2 data (both TACT and USGS) show less noise and image artefacts compared to the C1 data, due to the improved processing of the TIRS data during L1 generation. Example products are shown in Fig. 2 for Smartfin and Tidbit matchups, where the along track striping is rather obvious in the C1L1 data and much reduced in both C2 datasets. After TACT processing, i.e. with the same atmospheric correction parameters, the C2L1 data are generally a bit colder than the C1L1 data, as a result of the updated calibration in C2. A step change at around -117.5°E can be observed in the C2L2 product in the second example of the Smartfin

matchups (second row, last column of Fig. 2, 2017-08-05), here with a magnitude of about $0.5\text{--}1 \text{ K}$. A similar step change can be seen in the right hand side panel of the graphical abstract to this paper. These step changes are in fact present in all C2L2 scenes, and originate from the atmospheric parameters used to generate the ST, presumably as a result of a tiling grid used in the USGS atmospheric correction. TACT products use a per-pixel interpolated atmospheric correction which means these step changes are not present. C2L2 uses a lower B10 emissivity factor over water than TACT, which would partially explain higher ST for C2L2 compared with TACT.

3.2. Matchups

A total of 88 matchups were identified between in situ and L8, representing an in situ temperature range of about 20 K , between approximately 7 and 27°C . These matchups were distributed in the coastal zones of Hawaii (2), California (62), Costa Rica (2), UK (20) and Spain (2) (Fig. 3). There were 18 matchups with Tidbit data (May 2014–August 2018) and 70 matchups with Smartfin (June 2017–December 2019). Matchups were found throughout the year, with a higher number in the third quarter ($Q1 = 15$, $Q2 = 18$, $Q3 = 34$, $Q4 = 21$), corresponding to summer months in the northern hemisphere. Mean average temperatures (with standard deviation) per quarter were $Q1: 15.40 \pm 5.68^\circ\text{C}$, $Q2: 15.51 \pm 3.04^\circ\text{C}$, $Q3: 20.33 \pm 3.42^\circ\text{C}$, and $Q4: 17.16 \pm 3.82^\circ\text{C}$. The solar zenith angle ranged from 21° to 75° , with an average of 39.9° .

Scatter plots of the full 88 matchup dataset are presented in Fig. 4. These show that the scatter in the matchups is similar for the three processing methods, with uRMSD between 1.7 and 2.0 K . This is largely a result of the inherent sensor and image characteristics and the spatio-temporal mismatch between the in situ and satellite measurements. The MAD on the other hand shows that the bias increases from TACT C2L1 (0.8 K), to TACT C1L1 (1.3 K) and USGS C2L2 (1.7 K), with all methods on average slightly warmer compared to the in situ data.

In these matchups, relatively large vertical error bars are seen, that represent the IQR in the 5×5 pixel box. These are caused by spatial variability in the 5×5 satellite data box, largely as a result of close proximity to land (see e.g. Martí-Cardona et al., 2019). Points with large IQR show a larger warm bias of the satellite measurement. Especially the summer matchups from the San Diego area show large IQR due to the large difference in surface temperatures between land ($>50^\circ\text{C}$) and water (most points between $15\text{--}25^\circ\text{C}$).

Fig. 5 shows two horizontal transects through the matchup location for two of the San Diego area scenes (24 km ROI). These plots show the sharp transition between colder water and warmer land surface temperatures, and the position of the matchup location just at this transition, with the associated high IQR. For validation purposes it is justifiable to select stable targets, and hence matchups with high IQR ($> 0.25 \text{ K}$ in any of the processors) were removed. Additionally, 5×5 boxes where less than half of the pixels do not pass the optical data quality check were also removed, i.e. only matchups with $< 13/25$ MP using the OLI criteria listed above are retained. The MAD is relatively stable as function of MP, but the uRMSD increases rapidly when more than $12\text{--}15$ pixels are masked (not shown). The uRMSD increases rapidly for matchups where over half the pixels in the 5×5 pixel box are masked, i.e. matchups where the surfer is positioned at the land–water edge, and hence where half of the 5×5 pixel box contains land. This effect illustrates the difficulty of validating the nearshore ST, but also shows the advantage of using data logged by surfers with variable positions. The surfers are usually close to or at the land–water edge, but in about $1/3$ of our matchups their position is far enough to allow for a good validation matchup, while still being in the nearshore zone.

With this more stringent filtering on the satellite data, the matchup dataset reduces from 88 to 31 (5 Tidbit and 26 Smartfin), with about the same temperature range. The statistics for this reduced dataset improve considerably. Table 1 summarises the full and reduced matchup

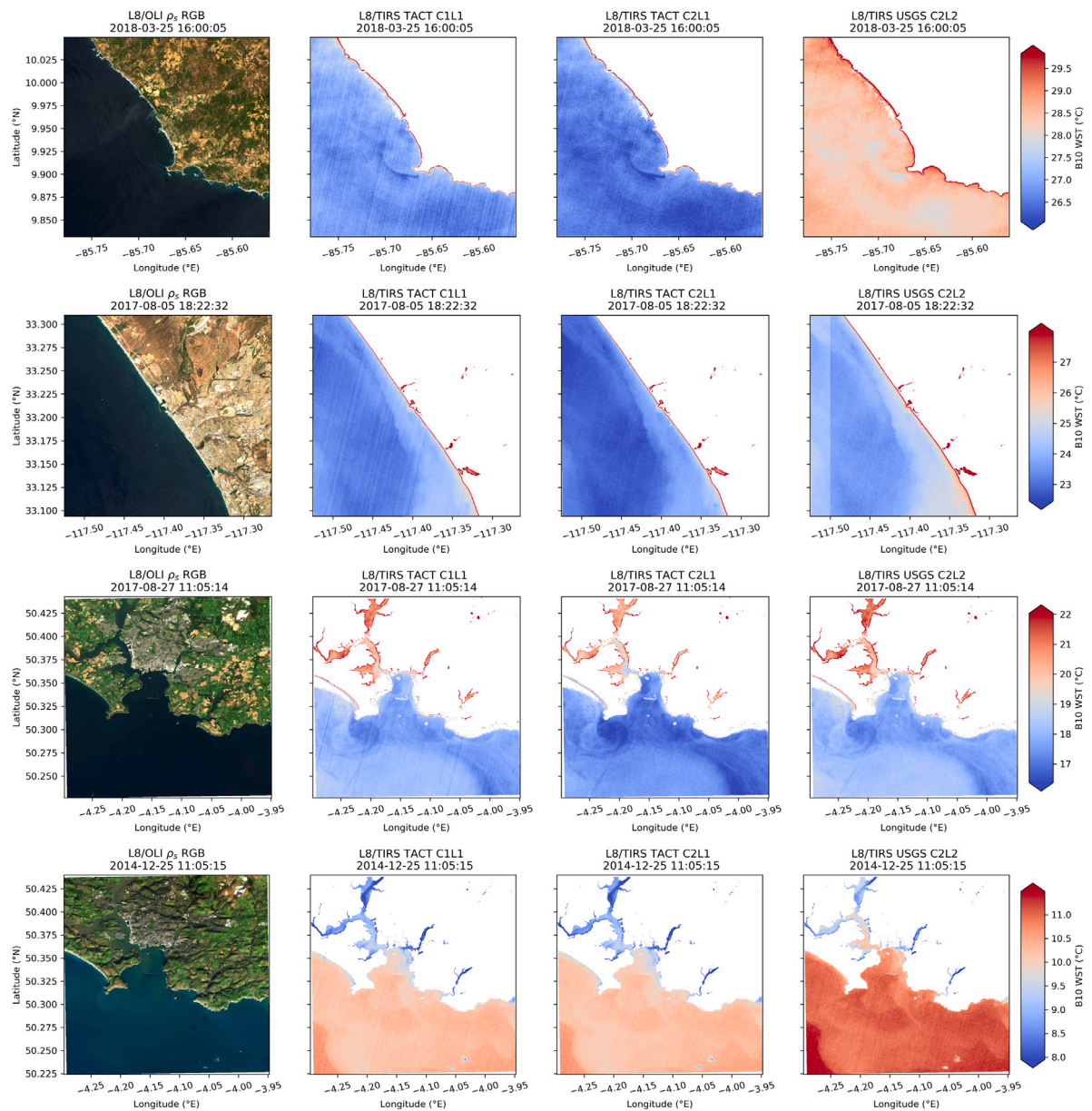


Fig. 2. Selected scenes for the matchups with Smartfin (top two rows, Costa Rica and California) and Tidbit (bottom two rows, UK summer and winter). Each row shows the OLI ρ_s RGB composite (655, 561, 483 nm reflectance between 0 and 0.15 linearly scaled to 8 bits) and the products derived from C1L1 and C2L1 data using TACT and the USGS C2L2 standard product. Non-water pixels are masked in white using the matchup filtering thresholds.

datasets. Fig. 6 shows the scatter plots of the high quality matchups, showing a reduction of the scatter to about 1.1 K uRMSD, and MAD ranging from 0.2 K (TACT C2L1), to 0.6 K (TACT C1L1) and 1.1 K (USGS C2L2). The TACT processed data has slightly higher biases (MAD <0.2 K versus <0.1 K) and higher noise (uRMSD 1.1 K versus 0.7 K) compared to previous results using Collection 1 data for offshore in situ measurements at towers and buoys (Vanhellemont, 2020a). This is likely the influence of the very nearshore location of the present matchups. The bias may be impacted by the bulk-to-skin correction applied to the in situ measurements, where we subtract 0.17 K from the in situ data to obtain equivalent skin temperature, which curiously is about the magnitude of the C2L1 TACT bias (0.19 K). Previous work has suggested that the skin and Smartfin differences may be smaller than 0.17 K due to the shallow depth of the Smartfin measurements (< 0.1 m), i.e. 0.13 K overall, and 0.06 K during the day, when the surfers would be collecting data (Brewin et al., 2021). Brewin et al. (2021) also suggested that in some cases, in a turbulent nearshore

environment, the skin cooling effect may disappear completely, which needs to be further investigated experimentally. RMA slopes for the matchup subset are 0.99 and 1.01 for TACT C1L1 and C2L1, while the slope of USGS C2L2 is > 1.07, indicating that for this product the bias increases with target temperature. On average the satellite data is biased warm, but the bias varies slightly with the target temperature. In Fig. 7 the difference between satellite and in situ measurement is shown as function of the target temperature. TACT shows a cold bias for most of the colder (<15 °C) and some of the warmer temperatures (>20 °C) encountered in the matchup dataset. USGS C2L2 only shows 4 points with negative biases, all for the colder (<15 °C) temperatures. For the higher temperatures (>20 °C) the bias of C2L2 is close to the dataset average bias. The rather triangular shape of the TACT results (Fig. 7 A and B) is caused by a remaining warm bias for some of the San Diego matchups in the 15–25 °C range, presumably as a result of large water–land temperature differences (Fig. 5). This effect is less visible in the USGS (Fig. 7 C) dataset, since their average bias also increases with

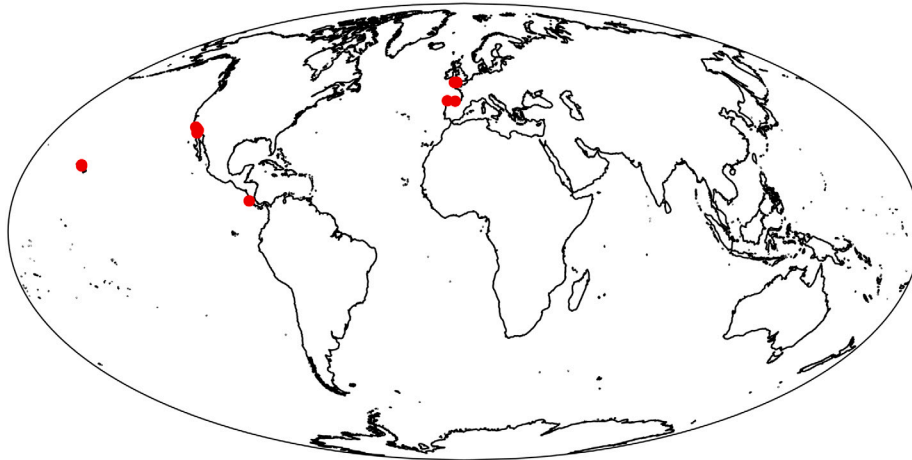


Fig. 3. Locations of the 88 total matchups. Tidbit data were available for California and UK coasts, while Smartfin data were obtained for all locations.

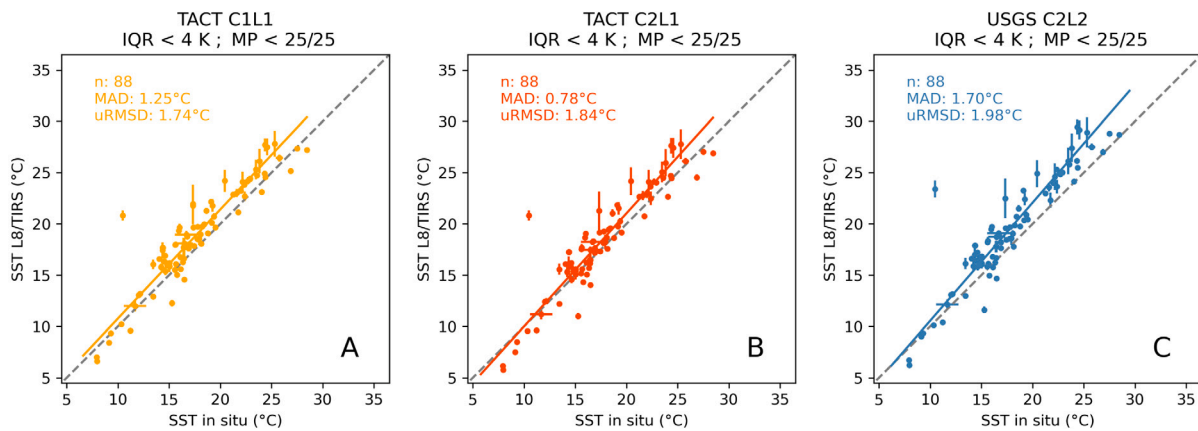


Fig. 4. Scatter plots for all 88 matchups for TACT C1L1 (A), TACT C2L1 (B), and USGS C2L2 (C). Error bars represent the IQR in a 5×5 pixel box (satellite) and the standard deviation over the surfing session.

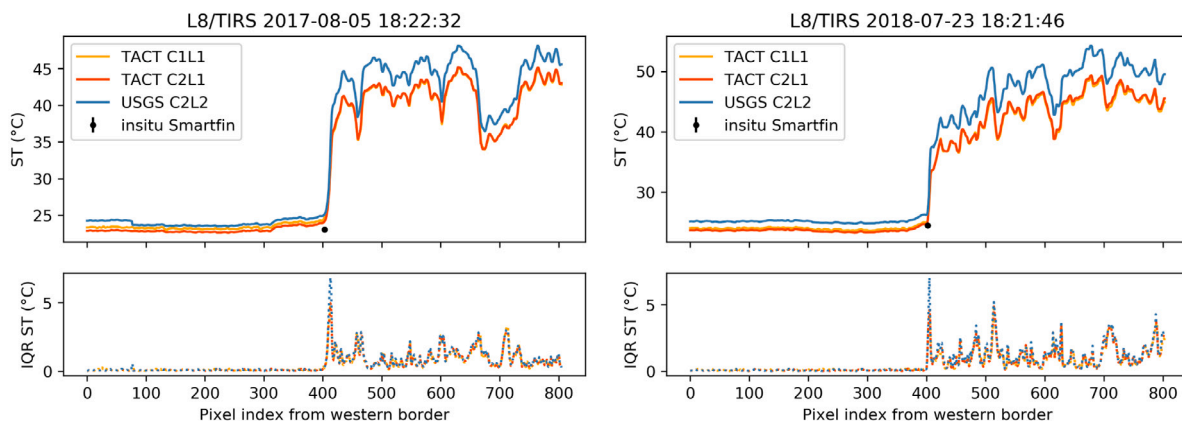


Fig. 5. Two transects through the 5×5 median filtered ST product for the San Diego coast, illustrating the large difference between water (pixels 0–400) and land (pixels 400–800) temperatures. The dotted lines represent the 5×5 filtered IQR. Note that TACT is used here with a fixed emissivity of water, and could hence underestimate the land temperatures. Note also in the left plot the step-change at around pixel 80 in the USGS C2L2 product as a result of tiling in the atmospheric correction.

target temperature. This bias increase is a result of the atmospheric parameters used in the USGS processing — see next paragraphs.

The differences between TACT C2L1 and USGS C2L2 data can be attributed to both the lower emissivity used by the latter, and by differences in the atmospheric parameters (L_u , L_d , and τ). When modifying the TACT processing to use the USGS C2L2 emissivity value, the MAD increases from 0.2 to 0.4 K (A panel of Fig. 8, and Table 1). Replacing

the atmospheric parameters by those of USGS C2L2, leads to an increase of MAD from 0.2 to 1.0 K (B panel of Fig. 8). The substitution of both emissivity and atmospheric parameters (essentially a local processing of the provided C2L1 to USGS C2L2) leads to an increase of > 1 K in MAD compared to TACT C2L1 processing (C panel of Fig. 8). The MAD of this product (1.2 K) is a bit higher than that of the USGS C2L2 (1.1 K) presumably due to the discretisation of the atmospheric parameters

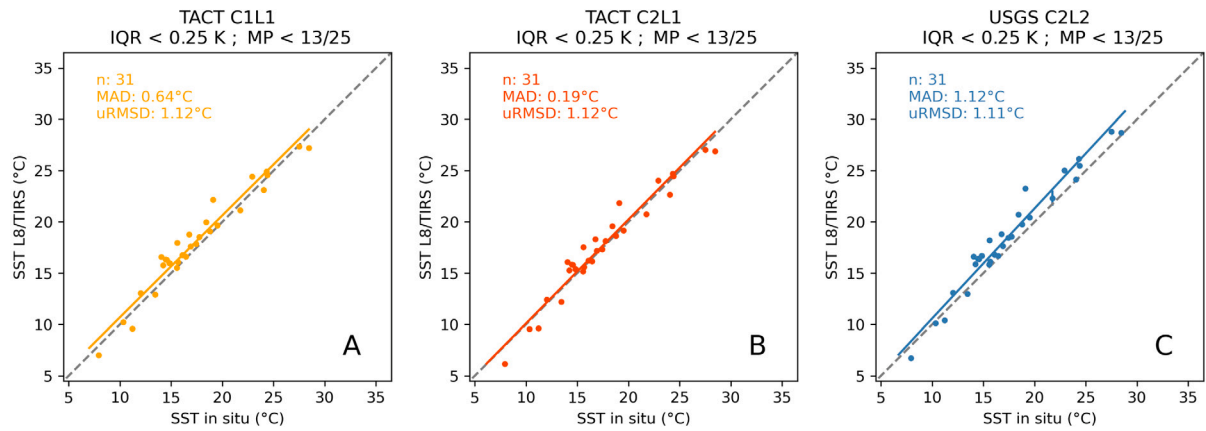


Fig. 6. Same as Fig. 4 but for the 31 matchups left after filtering on 5×5 box temperature stability ($\text{IQR} < 0.25$ K) and masking level ($< 13/25$ MP).

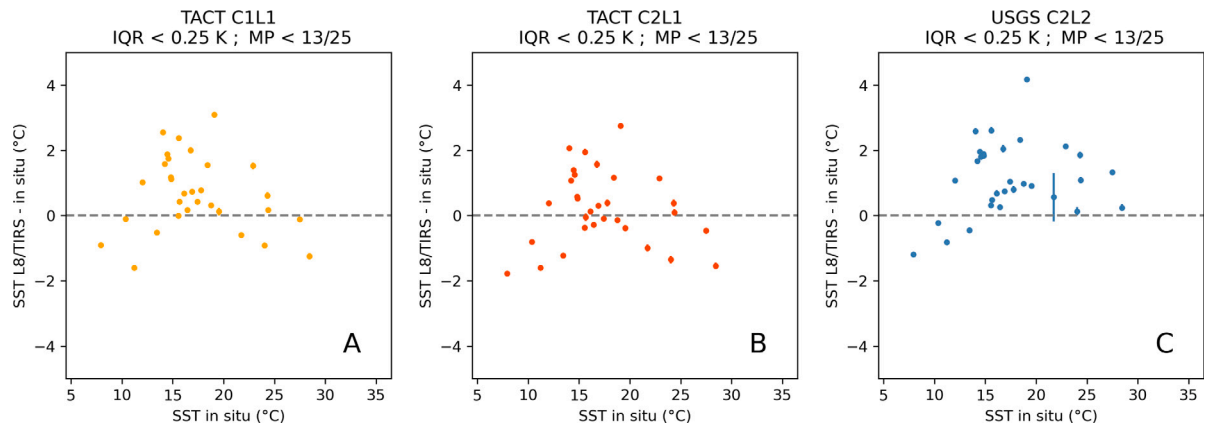


Fig. 7. Same as Fig. 6 but showing the difference between satellite and in situ measurement on the Y axis.

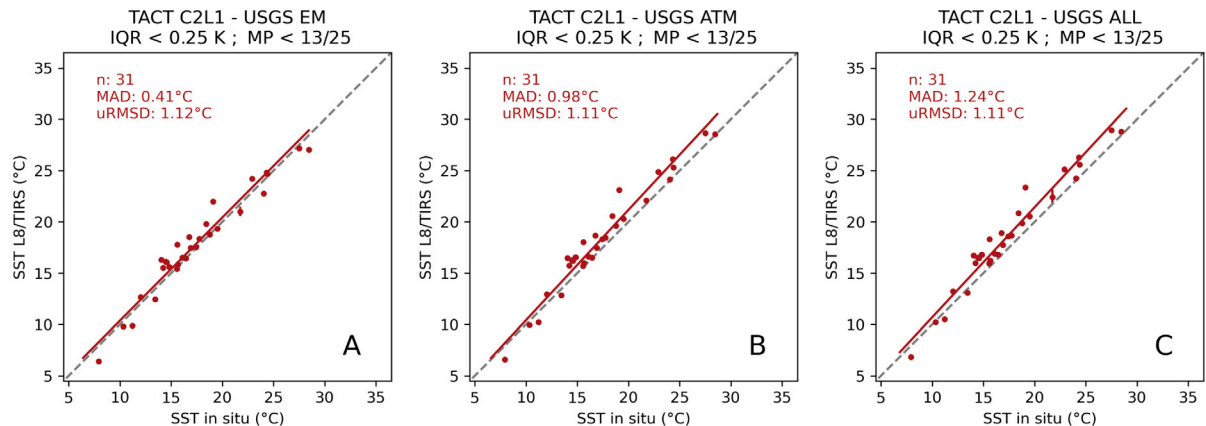


Fig. 8. Same as Fig. 6 but replacing TACT emissivity (left) atmospheric parameters (middle) and both (right) with those from USGS C2L2.

in the C2L2 GeoTIFF files. Atmospheric radiances are discretised to 1/1000, and the atmospheric transmittance and emissivity to 1/10000. The origin of the bias coming from the USGS atmospheric parameters is further analysed by comparing the retrieved L_d , L_u , and τ to those retrieved by TACT.

Fig. 9 shows the atmospheric parameters from USGS C2L2, and also those retrieved from the Atmospheric Correction Parameter Calculator (ACPC, Barsi et al. (2003, 2005), accessed 2021-03-20), compared to the ones derived in TACT for the full 88 matchups. The Ordinary Least Squared (OLS) regression line is used to estimate the relative differences between TACT and the other sources of atmospheric parameters.

A significant difference between TACT and USGS C2L2 atmospheric correction parameters is found for the L_d , where USGS values are about 30% of the TACT values. For L_d , ACPC retrieves an about 10% higher value compared to TACT. However, the impact of changes in L_d on the final ST product are minor due to the high emissivity of water (see Eq. (1)). The L_u and τ are very comparable for ACPC and TACT, with OLS slopes close to unity and low offsets (see also Vanhellemont, 2020a). Compared to TACT, for USGS C2L2 a lower τ (−6%) and for most of the range a higher L_u (offset +0.3 $\text{W m}^{-2} \text{sr}^{-1} \mu\text{m}^{-2}$ and slope 0.93) are found, both changes that lead to higher surface temperature retrievals. The underestimation of τ causes an additive increase of the

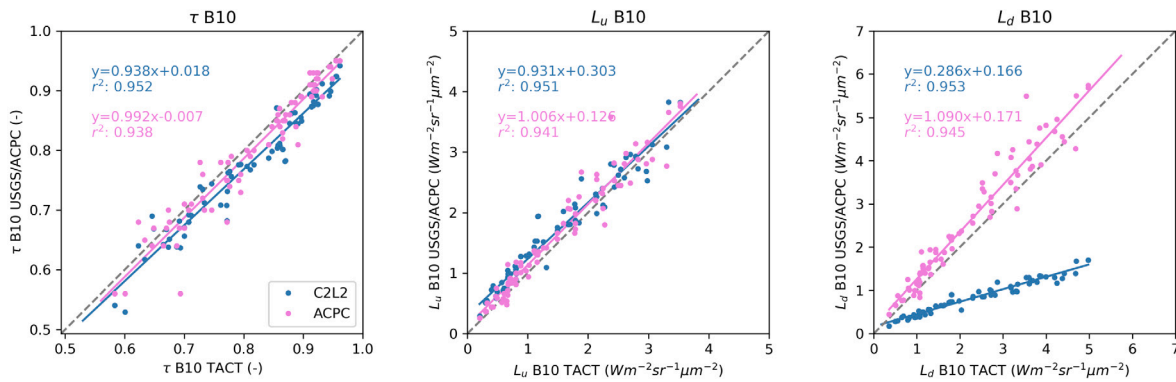


Fig. 9. A comparison of the atmospheric parameters retrieved for the 88 matchups. The Ordinary Least Squares regression line is shown as is the square of Pearson's r coefficient.

Table 1

Matchup performance of the different processors. n is the number of samples, and m the RMA regression slope. The matchup subsets are indicated by the number of matchups ($n = 88$ for all, and $n = 31$ for the highest quality matchups). Below the line results of changing the emissivity and atmospheric parameters in TACT to those from C2L2 are given.

Processor	Data Level	Emissivity	Atmosphere	n	MAD (K)	uRMSD (K)	m (1)
TACT	C1L1	TACT	TACT	88	1.253	1.745	1.064
TACT	C2L1	TACT	TACT	88	0.778	1.844	1.103
USGS	C2L2	USGS	USGS	88	1.703	1.979	1.152
TACT	C1L1	TACT	TACT	31	0.645	1.125	0.994
TACT	C2L1	TACT	TACT	31	0.194	1.122	1.014
USGS	C2L2	USGS	USGS	31	1.121	1.106	1.073
TACT	C2L1	USGS	TACT	88	1.005	1.916	1.096
TACT	C2L1	TACT	USGS	88	1.544	1.888	1.155
TACT	C2L1	USGS	USGS	88	1.819	1.981	1.153
TACT	C2L1	USGS	TACT	31	0.412	1.123	1.006
TACT	C2L1	TACT	USGS	31	0.977	1.108	1.076
TACT	C2L1	USGS	USGS	31	1.238	1.106	1.074

bias in ST, while the lower slope and non-zero offset for L_u causes a multiplicative increase of the bias in ST. In the scatter plots compared to in situ measurements, these effects are reflected in a vertical shift upwards and an increased slope of the USGS C2L2 points. The result of the slope in L_u is an increase of the bias as function of the target temperature. To summarise, the USGS C2L2 product has a higher warm bias compared to TACT, as a result of the lower emissivity value (causing around + 0.2 K in MAD), and the difference in atmospheric parameters (causing around + 0.8 K in MAD) used in C2L2 processing.

For a final comparison, matchups were also obtained using the single window calibration and code of Ermida et al. (2020). This code uses TCWV from the National Centers for Environmental Prediction (NCEP), and ASTER emissivity data, and can be run easily on the Collection 1 data hosted on Google Earth Engine (Gorelick et al., 2017). Valid results were obtained for 29 out of the 31 low-variability matchups, likely due to additional cloud filtering in the Ermida processing. For these matchups the statistics of the TACT and USGS processors are very similar to the ones obtained before, with the same scatter and slightly lower biases (Fig. 10). The Ermida code gives higher scatter (2 K) and higher biases (2 K) compared to the other processors, and hence is about 1 K warmer compared to USGS C2L2, and 1.5–2 K warmer compared to TACT. While its implementation in Google Earth Engine is convenient, this method cannot at present be recommended for nearshore water surface temperature retrieval.

The nearshore supports the highest levels of marine biodiversity and productivity in our ocean and is the region humans interact the most with. It is among the most dynamic regions on our planet, and requires observing systems capable of capturing this high spatial and temporal variability. High resolution satellite observations are a critical component of such observation systems, but to maximise their potential, they

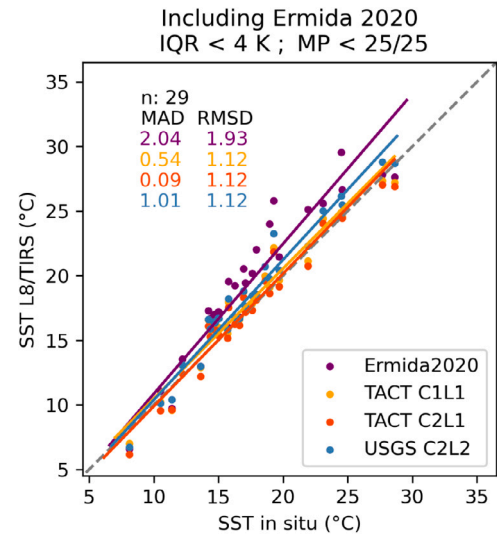


Fig. 10. Same as Fig. 8 but including results using the Google Earth Engine code of Ermida et al. (2020). 2 additional matchups where this code did not produce valid results were removed compared to Fig. 8.

need to be carefully compared with in situ measurements, to quantify accuracy and precision, essential for use in marine management applications. Owing to difficulties in deploying conventional oceanographic equipment in the nearshore, new solutions to in situ data collection in the nearshore are urgently needed. Here, we have demonstrated how one such citizen-science based solution (Brewin et al., 2017a), is capable of identifying the most accurate high-resolution satellite processor. Other in situ solutions exist to, including benthic temperature loggers (Brewin et al., 2018), coastal gliders (Rudnick et al., 2004), autonomous beach buoy systems (Shively et al., 2016) and the tagging of marine vertebrates with sensors (Fedak et al., 2004). Integrating all these observations, with satellite and models, into a wider nearshore observing system will ultimately help monitor and manage this critical region of our oceans, in the face of uncertain environmental change.

4. Conclusions

- Three surface temperature products derived from L8/TIRS were compared for retrieval of water surface temperatures: Collection 1 and Collection 2 top-of-atmosphere data as processed by TACT (TACT C1L1 and C2L1), and the USGS standard Collection 2 Level 2 product (USGS C2L2). In the most stable subset of matchup data, all three products showed a comparable uRMSD of 1.1 K, but for both C1 and C2 data, TACT gave lower biases (MAD respectively 0.6 and 0.2 K) compared to USGS C2L2 (1.1 K). An

additional processor implemented on a cloud computing platform was tested, but gave significantly larger bias and scatter.

- The higher bias of the USGS C2L2 product can be explained by differences in the emissivity value used over water (0.9880 for USGS C2L2 and 0.9926 for TACT), and differences in the estimated atmospheric parameters. The lower emissivity in the USGS C2L2 product adds around 0.2 K in MAD, and the difference in L_u and lower τ combined add about 0.8 K in MAD. The multiplicative error of the change in L_u results in a bias of C2L2 that increases with target temperature.
- The Collection 2 data shows less along-track striping in the output images compared to Collection 1 data. When Level 1 data is processed with TACT, Collection 2 shows lower biases compared to Collection 1 data. The standard Collection 2 Level 2 product shows a tiling grid in the atmospheric parameters and hence the output surface temperature, leading to step changes of about 0.5–1 K at the borders of this tiling grid. Overall, with these results Collection 2 data as processed by TACT show the most promising water surface temperature products from Landsat 8 with lowest biases and fewer processing artefacts.
- Data collected using low-cost contact thermometers integrated in sports equipment show promise for validation of high resolution satellite temperature products, but the transition zone between land and water has to be handled with care, especially when one of the two is significantly warmer than the other. Filtering based on the concurrent optical data from OLI, and the interquartile range of surface temperature in a 5×5 pixel box reduces the impact of land proximity and mixed pixels on the validation statistics. Surfing data could be similarly useful for validation of nearshore data from new missions such as L9/TIRS-2, launched in September 2021, and future missions, such as the Copernicus Land Surface Temperature Monitoring mission (LSTM), a High Priority Candidate Mission.

CRedit authorship contribution statement

Quinten Vanhellemont: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Visualization. **Robert J.W. Brewin:** Conceptualization, Methodology, Software, Data curation, Writing – original draft. **Philip J. Bresnahan:** Software, Data curation, Writing – review & editing. **Tyler Cyronak:** Software, Data curation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

TACT development was funded by the Federal Belgian Science Policy Office (BELSPO) under the BRAIN-be programme MICROBIAN project (BR/165/A1/MICROBIAN). Robert Brewin was supported by a UKRI Future Leader Fellowship (MR/V022792/1). The Lost Bird Foundation supported the development of the Smartfin project, including all Smartfin data acquisitions. The National Aeronautics and Space Administration (NASA) and the United States Geological Society (USGS) are thanked for acquiring and freely distributing Landsat 8 data. Google is thanked for hosting a copy of the Collection 1 Level 1 data on their Cloud Services. Smartfin is thanked for providing surfing data through their Application Platform Interface. The people recording and sharing data during their surfing sessions are thanked for their efforts, without which, this work would have not been possible. The authors of libRadtran are thanked for the release of their code under a GNU General Public License. The National Center for Atmospheric Research (NCAR) is thanked for hosting ERA5 data on a THREDDS/OPeNDAP server in their Research Data Archive (RDA). Two anonymous reviewers are thanked for their feedback on the paper.

References

- Anderson, M.C., Allen, R.G., Morse, A., Kustas, W.P., 2012. Use of landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sens. Environ.* 122, 50–65.
- Anderson, M., Gao, F., Knipper, K., Hain, C., Dulaney, W., Baldocchi, D., Eichmann, E., Hemes, K., Yang, Y., Medellin-Azuara, J., et al., 2018. Field-scale assessment of land and water use change over the California Delta using remote sensing. *Remote Sens.* 10 (6), 889.
- Augustine, J.A., DeLuisi, J.J., Long, C.N., 2000. SURFRAD—A National surface radiation budget network for atmospheric research. *Bull. Am. Meteorol. Soc.* 81 (10), 2341–2358.
- Barsi, J.A., Barker, J.L., Schott, J.R., 2003. An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument. In: IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477). 5, IEEE, pp. 3014–3016.
- Barsi, J., Schott, J., Hook, S., Raqueno, N., Markham, B., Radocinski, R., 2014. Landsat-8 thermal infrared sensor (TIRS) vicarious radiometric calibration. *Remote Sens.* 6 (11), 11607–11626.
- Barsi, J.A., Schott, J.R., Palluconi, F.D., Hook, S.J., 2005. Validation of a web-based atmospheric correction tool for single thermal band instruments. In: *Earth Observing Systems X*, Vol. 5882. International Society for Optics and Photonics, p. 58820E.
- Berk, A., Anderson, G.P., Bernstein, L.S., Acharya, P.K., Dothe, H., Matthew, M.W., Adler-Golden, S.M., Chetwynd Jr., J.H., Richtsmeier, S.C., Pukall, B., et al., 1999. MODTRAN4 Radiative transfer modeling for atmospheric correction. In: *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III*, Vol. 3756. International Society for Optics and Photonics, pp. 348–353.
- Brando, V., Braga, F., Zaggia, L., Giardino, C., Bresciani, M., Matta, E., Bellafiore, D., Ferrarin, C., Maicu, F., Benetazzo, A., Bonaldo, D., Falcieri, F., Coluccelli, A., Russo, A., S. C., 2015. High-resolution satellite turbidity and sea surface temperature observations of river plume interactions during a significant flood event. *Ocean Sci.* 11 (6), 909.
- Bresnahan, P.J., Cyronak, T., Martz, T., Andersson, A., Waters, S., Stern, A., Richard, J., Hammond, K., Griffin, J., Thompson, B., 2017. Engineering a smartfin for surf-zone oceanography. In: *OCEANS 2017-Anchorage*. IEEE, pp. 1–4.
- Brewin, R., Cyronak, T., Bresnahan, P., Andersson, A., Richard, J., Hammond, K., Billson, O., de Mora, L., Jackson, T., Smale, D., Dall'Olmo, G., 2020a. Sea surface temperature (SST) measurements collected by surfers around the coastlines of the southern UK, western Ireland, and San Diego, US, between 2014 and 2018. *British Oceanographic Data Centre—Natural Environment Research Council: Swindon, UK*.
- Brewin, R.J., Cyronak, T., Bresnahan, P.J., Andersson, A.J., Richard, J., Hammond, K., Billson, O., de Mora, L., Jackson, T., Smale, D., et al., 2020b. Comparison of two methods for measuring sea surface temperature when surfing. *Oceans* 1, 6–26.
- Brewin, R.J., de Mora, L., Billson, O., Jackson, T., Russell, P., Brewin, T.G., Shutler, J.D., Miller, P.I., Taylor, B.H., Smyth, T.J., et al., 2017b. Evaluating operational AVHRR sea surface temperature data at the coastline using surfers. *Estuar. Coast. Shelf Sci.* 196, 276–289.
- Brewin, R.J., de Mora, L., Jackson, T., Brewin, T.G., Shutler, J., 2015. On the potential of surfers to monitor environmental indicators in the coastal zone. *PLoS One* 10 (7), e0127706.
- Brewin, R.J., Hyder, K., Andersson, A.J., Billson, O., Bresnahan, P.J., Brewin, T.G., Cyronak, T., Dall'Olmo, G., de Mora, L., Graham, G., et al., 2017a. Expanding aquatic observations through recreation. *Front. Mar. Sci.* 4, 351.
- Brewin, R.J., Smale, D., Moore, P., Dall'Olmo, G., Miller, P., Taylor, B., Smyth, T., Fishwick, J., Yang, M., 2018. Evaluating operational AVHRR sea surface temperature data at the coastline using benthic temperature loggers. *Remote Sens.* 10 (6), 925.
- Brewin, R.J., Wimmer, W., Bresnahan, P.J., Cyronak, T., Andersson, A.J., Dall'Olmo, G., 2021. Comparison of a smartfin with an infrared sea surface temperature radiometer in the Atlantic Ocean. *Remote Sens.* 13 (5), 841.
- Cook, M.J., 2014. Atmospheric compensation for a landsat land surface temperature product.
- Cook, M., Schott, J., Mandel, J., Raqueno, N., 2014. Development of an operational calibration methodology for the landsat thermal data archive and initial testing of the atmospheric compensation component of a land surface temperature (LST) product from the archive. *Remote Sens.* 6 (11), 11244–11266.
- Davis, T.W., Berry, D.L., Boyer, G.L., Gobler, C.J., 2009. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of microcystis during cyanobacteria blooms. *Harmful Algae* 8 (5), 715–725.
- Donlon, C., Nightingale, T., Sheasby, T., Turner, J., Robinson, I., Emergy, W., 1999. Implications of the oceanic thermal skin temperature deviation at high wind speed. *Geophys. Res. Lett.* 26 (16), 2505–2508.
- Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T., et al., 2016. The libRadtran software package for radiative transfer calculations (version 2.0.1). *Geosci. Model Dev.* (5), 1647–1672.
- Ermiada, S.L., Soares, P., Mantas, V., Göttsche, F.-M., Trigo, I.F., 2020. Google earth engine open-source code for land surface temperature estimation from the landsat series. *Remote Sens.* 12 (9), 1471.

- Fedak, M., et al., 2004. Marine animals as platforms for oceanographic sampling: a "win/win" situation for biology and operational oceanography. *Memoirs Natl. Inst. Polar Res.* 58, 133–147.
- Freitas, S.C., Trigo, I.F., Macedo, J.a., Barroso, C., Silva, R., Perdigão, R., 2013. Land surface temperature from multiple geostationary satellites. *Int. J. Remote Sens.* 34 (9–10), 3051–3068.
- García-Santos, V., Cuxart, J., Martínez-Villagrasa, D., Jiménez, M., Simó, G., 2018. Comparison of three methods for estimating land surface temperature from landsat 8-tirs sensor data. *Remote Sens.* 10 (9), 1450.
- Gerace, A., Montanaro, M., 2017. Derivation and validation of the stray light correction algorithm for the thermal infrared sensor onboard landsat 8. *Remote Sens. Environ.* 191, 246–257.
- Giglio, L., Descloitres, J., Justice, C.O., Kaufman, Y.J., 2003. An enhanced contextual fire detection algorithm for MODIS. *Remote Sens. Environ.* 87 (2–3), 273–282.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27.
- Harley, C.D., Randall Hughes, A., Hultgren, K.M., Miner, B.G., Sorte, C.J., Thornber, C.S., Rodriguez, L.F., Tomanek, L., Williams, S.L., 2006. The impacts of climate change in coastal marine systems. *Ecol. Lett.* 9 (2), 228–241.
- Hulley, G.C., Hook, S.J., Abbott, E., Malakar, N., Islam, T., Abrams, M., 2015. The ASTER global emissivity dataset (ASTER GED): Mapping earth's emissivity at 100 meter spatial scale. *Geophys. Res. Lett.* 42 (19), 7966–7976.
- Ingleton, T., McMinn, A., 2012. Thermal plume effects: A multi-disciplinary approach for assessing effects of thermal pollution on estuaries using benthic diatoms and satellite imagery. *Estuar. Coast. Shelf Sci.* 99, 132–144.
- Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next landsat satellite: The landsat data continuity mission. *Remote Sens. Environ.* 122, 11–21.
- Jones, K., Gowen, R., 1990. Influence of stratification and irradiance regime on summer phytoplankton composition in coastal and shelf seas of the British Isles. *Estuar. Coast. Shelf Sci.* 30 (6), 557–567.
- Kaufman, Y.J., Justice, C.O., Flynn, L.P., Kendall, J.D., Prins, E.M., Giglio, L., Ward, D.E., Menzel, W.P., Setzer, A.W., 1998. Potential global fire monitoring from EOS-MODIS. *J. Geophys. Res.: Atmos.* 103 (D24), 32215–32238.
- Liu, L., Zhang, Y., 2011. Urban heat island analysis using the landsat TM data and ASTER data: A case study in Hong Kong. *Remote Sens.* 3 (7), 1535–1552.
- Malakar, N.K., Hulley, G.C., Hook, S.J., Laraby, K., Cook, M., Schott, J.R., 2018. An operational land surface temperature product for Landsat thermal data: Methodology and validation. *IEEE Trans. Geosci. Remote Sens.* 56 (10), 5717–5735.
- Martí-Cardona, B., Prats, J., Niclòs, R., 2019. Enhancing the retrieval of stream surface temperature from Landsat data. *Remote Sens. Environ.* 224, 182–191.
- Martínez, E., Antoine, D., d'Ortenzio, F., de Boyer Montégut, C., 2011. Phytoplankton spring and fall blooms in the North Atlantic in the 1980s and 2000s. *J. Geophys. Res. Oceans* 116 (C11).
- Merchant, C.J., Embury, O., Bulgin, C.E., Block, T., Corlett, G.K., Fiedler, E., Good, S.A., Mittaz, J., Rayner, N.A., Berry, D., et al., 2019. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data* 6 (1), 1–18.
- Merchant, C.J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C.E., Corlett, G.K., Good, S., McLaren, A., Rayner, N., Morak-Bozzo, S., et al., 2014. Sea surface temperature datasets for climate applications from phase 1 of the European space agency climate change initiative (SST CCI). *Geosci. Data J.* 1 (2), 179–191.
- Merchant, C.J., Harris, A., 1999. Toward the elimination of bias in satellite retrievals of sea surface temperature: 2. Comparison with in situ measurements. *J. Geophys. Res. Oceans* 104 (C10), 23579–23590.
- Minnett, P., Alvera-Azcárate, A., Chin, T., Corlett, G., Gentemann, C., Karagali, I., Li, X., Marsouin, A., Marullo, S., Maturi, E., et al., 2019. Half a century of satellite remote sensing of sea-surface temperature. *Remote Sens. Environ.* 233, 111366.
- Montanaro, M., Gerace, A., Lunsford, A., Reuter, D., 2014a. Stray light artifacts in imagery from the landsat 8 thermal infrared sensor. *Remote Sens.* 6 (11), 10435–10456.
- Montanaro, M., Levy, R., Markham, B., 2014b. On-orbit radiometric performance of the landsat 8 thermal infrared sensor. *Remote Sens.* 6 (12), 11753–11769.
- Murphy, R.R., Kemp, W.M., Ball, W.P., 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuar. Coasts* 34 (6), 1293–1309.
- Neukermans, G., Oziel, L., Babin, M., 2018. Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic. *Global Change Biol.* 24 (6), 2545–2553.
- Nogueira, M., Albergel, C., Boussetta, S., Johannsen, F., Trigo, I.F., Ermida, S.L., Martins, J.a., Dutra, E., 2020. Role of vegetation in representing land surface temperature in the CHTESSEL (CY45r1) and SURFEX-ISBA (v8. 1) land surface models: a case study over Iberia. *Geosci. Model Dev.* 13 (9), 3975–3993.
- O'Neil, J., Davis, T., Burford, M., Gobler, C., 2012. The rise of harmful cyanobacteria blooms: the potential roles of eutrophication and climate change. *Harmful Algae* 14, 313–334.
- Orth, R., Dutra, E., Trigo, I.F., Balsamo, G., 2017. Advancing land surface model development with satellite-based earth observations. *Hydrol. Earth Syst. Sci.* 21 (5), 2483–2495.
- Oziel, L., Neukermans, G., Ardyna, M., Lancelot, C., Tison, J.-L., Wassmann, P., Sirven, J., Ruiz-Pino, D., Gascard, J.-C., 2017. Role for Atlantic inflows and sea ice loss on shifting phytoplankton blooms in the Barents Sea. *J. Geophys. Res. Oceans* 122 (6), 5121–5139.
- Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C., Wang, W., 2002. An improved in situ and satellite SST analysis for climate. *J. Clim.* 15 (13), 1609–1625.
- Rudnick, D.L., Davis, R.E., Eriksen, C.C., Fratantoni, D.M., Perry, M.J., 2004. Underwater gliders for ocean research. *Mar. Technol. Soc. J.* 38 (2), 73–84.
- Sekertekin, A., 2019. Validation of physical radiative transfer equation-based land surface temperature using landsat 8 satellite imagery and SURFRAD in-situ measurements. *J. Atmos. Sol.-Terr. Phys.* 105161.
- Sekertekin, A., Bonaoni, S., 2020. Land surface temperature retrieval from landsat 5, 7, and 8 over rural areas: assessment of different retrieval algorithms and emissivity models and toolbox implementation. *Remote Sens.* 12 (2), 294.
- Senay, G.B., Friedrichs, M., Singh, R.K., Velpuri, N.M., 2016. Evaluating landsat 8 evapotranspiration for water use mapping in the Colorado River Basin. *Remote Sens. Environ.* 185, 171–185.
- Shively, D.A., Nevers, M.B., Breitenbach, C., Phanikumar, M.S., Przybyla-Kelly, K., Spoljaric, A.M., Whitman, R.L., 2016. Prototypic automated continuous recreational water quality monitoring of nine Chicago beaches. *J. Environ. Manag.* 166, 285–293.
- Trinh, R.C., Ficht, C.G., Gierach, M.M., Holt, B., Malakar, N.K., Hulley, G., Smith, J., 2017. Application of landsat 8 for monitoring impacts of wastewater discharge on coastal water quality. *Front. Mari. Sci.* 4, 329.
- Trombetta, T., Vidussi, F., Mas, S., Parin, D., Simier, M., Mostajir, B., 2019. Water temperature drives phytoplankton blooms in coastal waters. *PLoS One* 14 (4), e0214933.
- USGS, 2020. Collection 2 landsat 8-9 OLI (operational land imager) and TIRS (thermal infrared sensor) level-2 science product. <http://dx.doi.org/10.5066/P9OGBGM6>.
- Vanhellemont, Q., 2019. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. *Remote Sens. Environ.* 225, 175–192. <http://dx.doi.org/10.1016/j.rse.2019.03.010>.
- Vanhellemont, Q., 2020a. Automated water surface temperature retrieval from landsat 8/TIRS. *Remote Sens. Environ.* 237, 111518. <http://dx.doi.org/10.1016/j.rse.2019.03.010>.
- Vanhellemont, Q., 2020b. Combined land surface emissivity and temperature estimation from landsat 8 OLI and TIRS. *ISPRS J. Photogramm. Remote Sens.* 166, 390–402.
- Vanhellemont, Q., 2020c. Sensitivity analysis of the dark spectrum fitting atmospheric correction for metre- and decametre-scale satellite imagery using autonomous hyperspectral radiometry. *Opt. Express* 27 (20), A1372–A1399. <http://dx.doi.org/10.1364/OE.397456>.
- Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* 34 (4), 892–905.
- Wan, Z., Wang, P., Li, X., 2004. Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern great plains, USA. *Int. J. Remote Sens.* 25 (1), 61–72.
- Wilhelm, S., Adrian, R., 2008. Impact of summer warming on the thermal characteristics of a polymictic lake and consequences for oxygen, nutrients and phytoplankton. *Freshwater Biol.* 53 (2), 226–237.
- Wong, P.P., Losada, I.n.J., Gattuso, J.-P., Hinkel, J., Khattabi, A., McInnes, K.L., Saito, Y., Sallenger, A., et al., 2014. Coastal systems and low-lying areas. *Clim. Chang.* 2104, 361–409.
- Yuan, F., Bauer, M.E., 2007. Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in landsat imagery. *Remote Sens. Environ.* 106 (3), 375–386.